Impact of artificial intelligence on cybersecurity: emerging threats and preventive measures

Jairo Michael Rojas Hilario, BEng[1], Aureliano Sanchez García, MS[2], and Jhon Jhonathan Peñalva Sanchez, PhD[3]

[1]*Universidad Tecnológica del Perú, Perú, U18102746@utp.edu.pe*
[2]*Universidad Tecnológica del Perú, Perú, C26533@utp.edu.pe*
[3]*Universidad Tecnológica del Perú, Perú, C25434@utp.edu.pe*

Abstract–The rise of artificial intelligence (AI) has accelerated technological development, transforming sectors like cybersecurity. In this field, AI has become a key tool for detecting and mitigating cyber threats, but it is also being used to create more sophisticated and difficult-to-persuade attacks. This duality presents a significant challenge. On one hand, AI enhances defenses against vulnerabilities; on the other, cybercriminals use it to develop more complex attacks. The objective of this review is to analyze these new threats and the methods for detecting and preventing them. To do this, the PICO methodology will be used for a focused literature search, complemented by the PRISMA method for the selection and analysis of articles. The integration of AI in cybersecurity seeks to reduce the frequency and severity of attacks by improving the efficiency of threat detection. While this technology offers significant improvements, it is crucial to conduct a thorough analysis before its implementation. Companies must adapt and train their teams to effectively integrate AI into their security systems, thereby strengthening their defenses against the growing landscape of cyber threats.

Keywords-- Artificial Intelligence, cybersecurity, vulnerability, methods, and prevention.

Impacto de la inteligencia artificial en la ciberseguridad: amenazas emergentes y medidas preventivas

Jairo Michael Rojas Hilario, BEng[1], Aureliano Sanchez García, MS[2], and Jhon Jhonathan Peñalva Sanchez, PhD[3]

[1]*Universidad Tecnológica del Perú, Perú, U18102746@utp.edu.pe*
[2]*Universidad Tecnológica del Perú, Perú, C26533@utp.edu.pe*
[3]*Universidad Tecnológica del Perú, Perú, C25434@utp.edu.pe*

Resumen–El auge de la inteligencia artificial (IA) ha acelerado el desarrollo tecnológico, transformando sectores como el de la ciberseguridad. En este ámbito, la IA se ha convertido en una herramienta clave para detectar y mitigar las amenazas cibernéticas, pero también se está utilizando para crear ataques más sofisticados y difíciles de persuadir. Esta dualidad presenta un desafío significativo. Por un lado, la IA potencia las defensas contra vulnerabilidades; por otro, los ciberdelincuentes la emplean para desarrollar ataques más complejos. El objetivo de esta revisión es analizar estas nuevas amenazas y los métodos para detectarlas y prevenirlas. Para ello, se utilizará la metodología PICO para una búsqueda bibliográfica focalizada, complementada con el método PRISMA para la selección y el análisis de artículos. La integración de la IA en la ciberseguridad busca reducir la frecuencia y gravedad de los ataques al mejorar la eficiencia en la detección de amenazas. Si bien esta tecnología ofrece mejoras significativas, es crucial realizar un análisis exhaustivo antes de su implementación. Las empresas deben adaptarse y capacitar a sus equipos para integrar la IA de manera efectiva en sus sistemas de seguridad, fortaleciendo así sus defensas contra el creciente panorama de amenazas cibernéticas.

Palabras clave-- Inteligencia Artificial, ciberseguridad, vulnerabilidad, métodos y prevención.

I. Introducción

En los últimos 10 años, muchos sectores fueron avanzando de forma acelerada con la inteligencia artificial, entre ellos, la ciberseguridad. Teniendo la habilidad de manejar grandes cantidades de información, identificar patrones complejos y tomar decisiones de forma autónoma, permitiendo desarrollar formas de defensa más eficientes ante los ataques cibernéticos. Sin embargo, también usan esta tecnología para diseñar nuevas amenazas más sofisticados.

Como resultado de lo anterior, el uso de la IA tiene dos maneras de expandirse en el ámbito de la ciberseguridad. Por un lado, desarrollar mecanismos de detección y prevención de ataques; y por el otro lado, la creación de amenazas que utilizan capacidades complejas de IA para provocar ciberataques intrincados y complicados [1].

Por ese motivo, se necesita realizar una revisión sistemática que permita identificar las principales amenazas que están desarrolladas con IA, así como también las estrategias propuestas que ayuden a contrarrestarlas. Un análisis así nos ayudaría a comprender el impacto que tiene la IA en la seguridad digital, donde la tecnología avanza, como también sus riesgos.

Para garantizar que esta investigación contenga información relevante, se utilizará la metodología PICO para describir de manera precisa los componentes del tema. Esto permitirá realizar una búsqueda clara y precisa en la base de datos Scopus, con el fin de recolectar información de alta calidad. La elección de PICO se debe a que es la metodología más utilizada para la formulación clara de preguntas, ya que facilita la definición del problema en la búsqueda de evidencia y proporciona un marco común para evaluar investigaciones científicas [2]. Además, se implementará el método PRISMA, una herramienta diseñada para optimizar la calidad y transparencia de las revisiones sistemáticas. Este método proporciona directrices rigurosas para la identificación, selección, evaluación y síntesis de la evidencia [2]. Por lo tanto, PRISMA será fundamental para elegir y evaluar artículos de calidad que estén directamente relacionados con la inteligencia artificial y la ciberseguridad. El aporte de esta investigación es que realiza un análisis sistemático sobre la dualidad de la inteligencia artificial (IA) en la ciberseguridad. El estudio demuestra cómo la IA se utiliza para crear ataques más sofisticados, mientras que, a su vez, funciona como una herramienta clave para la detección y mitigación efectiva de amenazas, proporcionando una base sólida para el desarrollo de futuras soluciones en el campo.

El presente artículo se estructura de la siguiente manera: La introducción presenta el contexto y la problemática. Posteriormente, la sección de Metodología detalla el diseño de la investigación, el uso del método PICO para la formulación de preguntas, las bases de datos utilizadas (específicamente Scopus), el método PRISMA para la selección de artículos, y el proceso de búsqueda y selección de estudios. A continuación, se presentan los Resultados obtenidos. Finalmente, las secciones de Discusión y Conclusiones analizan los hallazgos, sus implicaciones y las futuras líneas de investigación.

II. Metodología

A. Diseño de la investigación

En el diseño de la investigación, se detalla el procedimiento para la obtención de artículos que analizan las amenazas y los métodos de mitigación con inteligencia artificial (IA) en ciberseguridad, con especial énfasis en cómo los sistemas tradicionales se vuelven vulnerables. Para ello, se empleará el marco estratégico PICO para formular una estrategia de búsqueda precisa, basándose en sus cuatro componentes: Población, Intervención, Comparación y Resultado. Adicionalmente, se aplicará el método PRISMA para un análisis exhaustivo y una selección rigurosa de los artículos obtenidos.

B. Método PICO

Para formular la pregunta de investigación, se utilizaron los cuatro componentes del marco PICO, como se detalla en la Tabla I. La Población se centra en los sistemas tradicionales de ciberseguridad. La Intervención consiste en la implementación de la inteligencia artificial (IA) para la detección y prevención de incidentes. La Comparación evalúa la efectividad de las metodologías tradicionales frente a la implementación de la IA. Finalmente, el Resultado esperado es que el uso de la IA mejore significativamente la eficiencia en la detección de amenazas, fortaleciendo así la seguridad de los sistemas.

TABLA I
DESARROLLO DE LA COMPONENTE PICO EN FUNCIÓN A LA INVESTIGACIÓN

Componente	Descripción
P	Las empresas de ciberseguridad enfrentan constantes ataques automatizados, impulsados por tecnologías avanzadas de inteligencia artificial, que explotan vulnerabilidades en sistemas tradicionales.
I	El manejo de modelos de IA para la detección y prevención de amenazas cibernéticas refuerza la capacidad de actuar frente a las amenazas de seguridad.
C	Comparado con las metodologías tradicionales de ciberseguridad, el uso de IA ofrece una mayor capacidad para identificar patrones y amenazas emergentes en tiempo real.
O	Con el uso de la inteligencia artificial en ciberseguridad se aumentaría el análisis para detectar amenazas, teniendo resultados a través de su efectividad y mejorando la prevención de ataques cibernéticos, asi teniendo una mejora en la seguridad.

Posteriormente, en la Tabla II, y basándose en el marco PICO, se crearon preguntas específicas para cada componente (QR1, QR2, QR3, QR4). Este proceso derivó en la pregunta orientadora principal (QR) del estudio: ¿De qué manera el uso de la inteligencia artificial afecta la detección, prevención y respuesta ante amenazas en las empresas de ciberseguridad, en comparación con los métodos tradicionales?

TABLA II
DESARROLLO DE LAS PREGUNTAS PICO EN FUNCIÓN A LA INVESTIGACIÓN

QR	¿Las empresas de ciberseguridad, de qué manera afecta el uso de inteligencia artificial orientada a la detección, prevención y respuesta ante amenazas a comparación con los métodos tradicionales de ciberseguridad?
QR1	¿Cómo afectan los ataques automatizados impulsados por tecnologías avanzadas de IA a la ciberseguridad, en cuanto a las vulnerabilidades de los sistemas tradicionales?
QR2	¿De qué manera la inteligencia artificial mejora la detección de amenazas y prevención de vulneraciones, así como la capacidad de respuesta a los ataques?
QR3	¿Qué tan efectivo son las metodologías tradicionales comparado con el uso de IA para la identificación de patrones y la mitigación de amenazas emergentes?
QR4	¿Cuál es el efecto de la inteligencia artificial en la tasa de detección, la reducción de tiempo para mitigar ataques y la mejora en la seguridad?

Finalmente, como último paso, se identifican las palabras clave para cada componente del PICO, tal como se muestra en la Tabla III. Esto es crucial para formular la ecuación de búsqueda que nos permitirá encontrar artículos precisos y relevantes para nuestra investigación. Para este estudio, se ha elegido la base de datos Scopus, ya que se destaca por su gran cantidad de métricas de investigación en comparación con otras bases de datos [3].

TABLA III
DESCRIPCIÓN DE LAS PALABRAS CLAVES DEL MÉTODO PICO

Componentes	Palabras claves (Keywords)
P (population)	cybersecurity companies, automated attacks, advanced technologies, artificial intelligence, and vulnerabilities in systems
I (intervention)	artificial intelligence, detection, prevention, response capability, security incidents
C (comparison)	traditional methodologies, cybersecurity, use of AI, identify patterns and emerging threats
O (outcome)	artificial intelligence, threat detection, effectiveness, prevention of cyber-attacks and improvement in security

C. Base de datos Scopus

Una vez que se obtiene la ecuación de búsqueda, se introduce en la base de datos de Scopus. Para ello, se utilizan los conectores booleanos 'OR' para buscar cada una de las palabras clave de un mismo componente y 'AND' para combinar los diferentes componentes de la ecuación. Esta estrategia permite obtener una mayor precisión al momento de seleccionar los artículos más relevantes para el tema de investigación: *(TITLE-ABS-KEY ("cybersecurity companies"*

OR "automated attacks" OR "advanced technologies" OR "artificial intelligence" OR "Vulnerabilities in systems") AND TITLE-ABS-KEY ("artificial intelligence" OR detection OR prevention OR responsiveness OR "security incidents") AND TITLE-ABS-KEY ("traditional methodologies" OR cybersecurity OR "use of AI" OR "identify patterns" OR "emerging threats") AND TITLE-ABS-KEY ("artificial intelligence" OR "threat detection" OR effectiveness OR "cyberattack prevention" OR "improved security"))

A continuación, en la Tabla IV, se presentan los criterios de inclusión y exclusión. Esta organización nos servirá como una guía clara para determinar qué características deben tener los artículos que seleccionemos para nuestra investigación, asegurando así su relevancia y calidad.

TABLA IV
CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN

Criterios de Inclusión	Criterios de Exclusión
CI1: Deben tener relación con La adopción de la IA en el ámbito de la ciberseguridad.	CE1: Deben hablar de como idea principal sobre la ciberseguridad Y EL área de la informática.
CI2: Los artículos deben tener información precisa sobre las amenazas que fueron diseñados con IA, que exploten las vulnerabilidades.	CE2: Los artículos DEBEN HABLAR sobre las amenazas y formas de mitigar a los sistemas tradicionales.
CI3: Deben tambien tener relación a medidas de cómo prevenir y de cómo mejorar la detección de estas amenazas, explicando como lo haría.	CE3: Solo artículos que sean de acceso abierto y que sean publicados en el rango del 2020 y 2025.

D. Método PRISMA

Al implementar la fórmula PICO en la base de datos de Scopus, se obtuvieron inicialmente 13,523 documentos. Posteriormente, se aplicaron los filtros de Scopus para refinar la búsqueda, lo que redujo la cantidad a 1,136 artículos. Tras un segundo filtrado más detallado, se seleccionaron 354 documentos por su relevancia. A continuación, se descartaron 100 artículos que estaban incompletos o no se podían visualizar. Finalmente, se realizó una revisión exhaustiva para asegurar que los documentos cumplieran con los criterios de inclusión, eliminando duplicados y aquellos que no eran directamente relevantes. Este proceso culminó con la selección de 38 artículos para la Revisión Sistemática de la Literatura, como se muestra en la Fig.1

E. Proceso de búsqueda y selección de estudios

Para garantizar la calidad de los artículos seleccionados, se aplicó la metodología PRISMA 2020, que consta de cuatro etapas principales:

1) Identificación: Se inició la búsqueda exclusivamente en la base de datos Scopus, donde se obtuvieron un total de 13,523 documentos. Mediante el uso de filtros, se refinó la búsqueda para incluir solo artículos publicados entre 2020 y 2025 en el área de la informática. Además, se limitó el tipo de documento a "artículos" y se incluyeron las palabras clave "Inteligencia Artificial", "Ciberseguridad" y "Seguridad de red". Finalmente, se filtraron para que fueran de acceso abierto, resultando en 1,136 artículos para la siguiente etapa.

2) Cibrado: En esta fase, los 1,136 artículos fueron revisados exhaustivamente. Se evaluó el título y el resumen de cada uno para asegurar que su contexto y criterios fueran directamente relevantes para la investigación, lo que redujo la selección a 354 artículos.

3) Elegibilidad: Aquí se verificó la accesibilidad y la integridad de los documentos. Se descartaron 100 artículos que estaban incompletos o no eran accesibles, dejando un total de 254. Posteriormente, se realizó una revisión más minuciosa para confirmar que cumplieran con todos los criterios de inclusión. En este paso se eliminaron los artículos duplicados y aquellos que no eran lo suficientemente relevantes, lo que dejó una selección final de 38 artículos.

4) Incluidos: Después de completar el análisis detallado a través de las etapas anteriores, se obtuvieron 38 artículos que serán utilizados para la presente investigación.

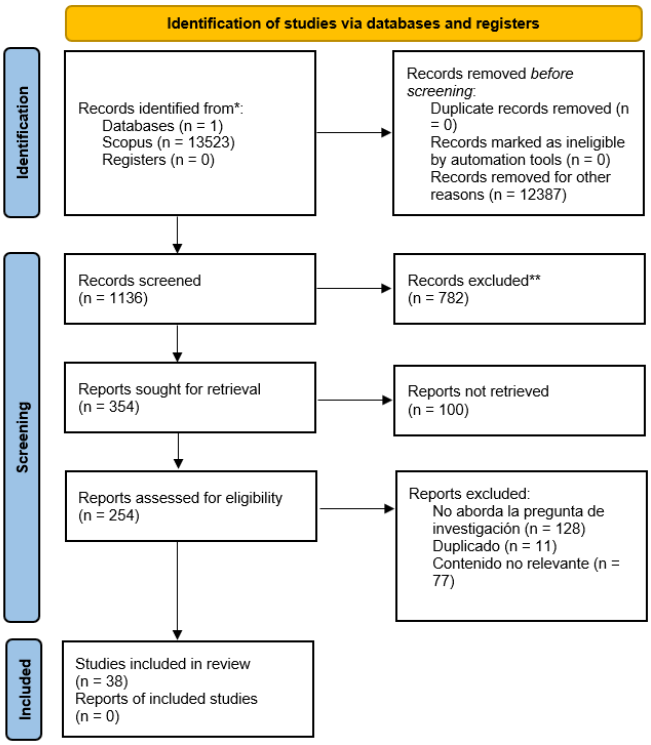

Fig. 1 Diagrama de flujo PRISMA

III. RESULTADOS

A partir de los 38 artículos seleccionados con las metodologías PICO-PRISMA, se obtuvieron los siguientes resultados. La Fig. 2 muestra la distribución de artículos por país. Se observa que Arabia Saudita y Estados Unidos lideran la lista con la misma cantidad, 6 artículos cada uno. Les siguen

Egipto y Turquía, ambos con 2 artículos, y el resto de países con una menor cantidad.

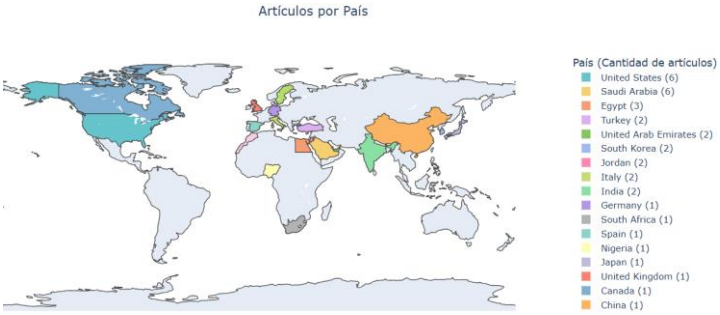

Fig. 2 Grafica de la cantidad de artículos obtenidos por país

En la Fig. 3, se presenta la distribución de los 38 artículos seleccionados por año, abarcando el periodo de 2020 a 2025. Los resultados muestran que la mayoría de los estudios son muy recientes: 18 artículos corresponden al año 2025 y 16 al 2024. Por otro lado, se encontraron 3 artículos del año 2023 y 1 del 2022. No se seleccionaron artículos de los años 2021 y 2020, lo que subraya la actualidad y el enfoque reciente de esta investigación. En la Fig. 4 se puede visualizar las 15 palabras claves que aparecen con más frecuencia en los artículos siendo entre las principales palabras como Ciberseguridad, Machine Learning, Deep Learning, Inteligencia Artificial, entre otros.

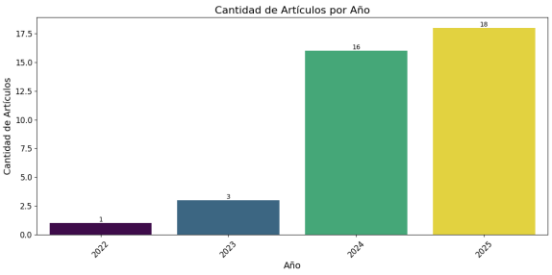

Fig. 3 Cantidad de artículos obtenidos por año

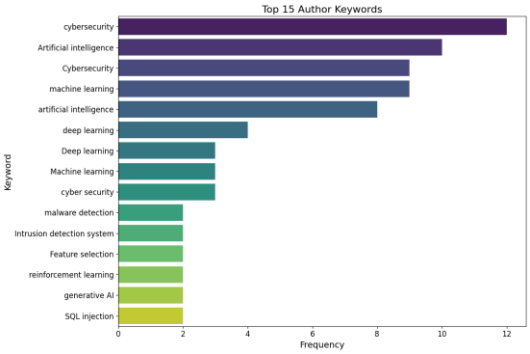

Fig. 4 Palabras claves más frecuentes

A. ¿Cómo afectan los ataques automatizados impulsados por tecnologías avanzadas de IA a la ciberseguridad, en cuanto a las vulnerabilidades de los sistemas tradicionales?

Los ataques automatizados impulsados por tecnologías avanzadas de inteligencia artificial (IA) tienen un impacto considerable en la ciberseguridad, ya que aprovechan las vulnerabilidades de los sistemas tradicionales. Los sistemas de detección de intrusiones (IDS) convencionales, por ejemplo, carecen de la capacidad para diferenciar entre el tráfico legítimo y el malicioso, lo que permite que los ataques pasen desapercibidos. Una de las amenazas emergentes es la ofuscación de ataques, que utiliza redes generativas antagónicas (GANs) para generar tráfico de red que simula ser real, ocultando así los ataques de los IDS tradicionales [4]. Otra táctica consiste en que los atacantes utilicen servicios legítimos como Dropbox para enviar comandos y control, evadiendo la detección de los sistemas de seguridad [5]. Además, el uso de grandes modelos de lenguaje (LLMs) en el phishing lateral permite la creación de correos electrónicos más sofisticados y difíciles de identificar por los métodos convencionales [6]. La creación automática de ataques también se beneficia de LLMs para generar código malicioso que es cada vez más difícil de detectar [7]. Asimismo, herramientas como ChatGPT se utilizan para suplantar identidades, facilitando ataques más personalizados que los sistemas tradicionales no pueden identificar [8]. Por último, los ataques dinámicos basados en el aprendizaje por refuerzo emplean técnicas que se adaptan y evaden las defensas fijas basadas en firmas [9, 10]. En conclusión, los avances en IA están superando las capacidades de detección de los sistemas tradicionales, lo que hace que los sistemas desprevenidos sean vulnerables a estos nuevos tipos de amenazas. La Tabla V resume las principales vulnerabilidades que los ataques automatizados aprovechan en los sistemas tradicionales.

B. ¿De qué manera la inteligencia artificial mejora la detección de amenazas y prevención de vulneraciones, así como la capacidad de respuesta a los ataques?

La inteligencia artificial (IA) optimiza de manera considerable la identificación de amenazas, la prevención de brechas de seguridad y la capacidad de respuesta ante diversos tipos de ataques, a través del uso de enfoques y modelos avanzados. A continuación, se detalla cómo los métodos impulsados por IA pueden ayudar a detectar y mitigar amenazas según el tipo de ataque:

TABLA V
VULNERABILIDADES EN MÉTODOS TRADICIONALES ANTE ATAQUES IMPULSADOS POR IA

Ataques impulsados por IA	Definición	Vulnerabilidades en Métodos Tradicionales	Referencia
Ofuscación de ataques	Uso de redes GANs para generar tráfico de red para ocultar ataques reales	Los IDS tradicionales no son capaces de distinguir entre el tráfico real y el malicioso haciendo que los ataques pasen desapercibidos	[4]
Uso de servicios legítimos para CnC	Los atacantes se comunican con sistemas comprometidos con servicios legítimos con Dropbox, para evitar ser detectados	Los sistemas tradicionales no están configurados para detectar este tipo de tráfico, donde están comunicado el malware y servidores CnC que usan servicios legítimos	[5]
Phishing lateral con LLMs	Uso de modelos de lenguaje grande (LLMs) para generar correos de phishing más sofisticados y personalizados	Los sistemas que detectan phishing de forma tradicional, mediante la identificación de errores de tipeado, no logran detectar correos generados por IA que son más convincentes y sin errores ortográficos	[6]
Generación automatizada de técnicas de ataque	Usando LLMs para generar código de ataque que va desde los vectores iniciales hasta las acciones relacionadas con el impacto, todo de forma automatizada.	Los sistemas tradicionales no están preparados para detectar ataques generados automáticamente, ya que la IA puede crear código sofisticado que elude las defensas basadas en firmas y patrones predefinidos.	[7]
Phishing y suplantación de identidad con ChatGPT	El uso de ChatGPT para generar correos electrónicos de phishing, páginas falsas de inicio de sesión y malware, realizando ataques más sofisticados	Los sistemas tradicionales de detección de phishing no están preparados para ataques personalizados generados por IA y son más difíciles de detectar	[8]
Ataques automatizados y dinámicos	Se utiliza aprendizaje por refuerzo (RL) para automatizar la creación de ataques sofisticados, como phishing, explotación de vulnerabilidades y exfiltración de datos mediante técnicas adaptativas y personalizadas	Los sistemas tradicionales basados en firmas estáticas no pueden identificar ataques personalizados creados por IA, lo que permite a los atacantes evadir las defensas utilizando tácticas nuevas y adaptativas	[9, 10]

1. Dominio Malicioso: Estos ataques se centran en la utilización de sitios web para realizar phishing, distribuir malware o controlar bots, todo diseñado para evitar la detección. Para abordar este tipo de amenazas, se utiliza una Red Neuronal Superficial (SNN), mejorando la precisión y rendimiento de la detección en tiempo real [11]

2. Ataques en IoT (Intrusión): En los ataques a redes de dispositivos IoT, los atacantes buscan explotar vulnerabilidades en dispositivos conectados como cámaras de seguridad o termostatos. Para detectar y responder a estos ciberataques de manera efectiva, se emplea IA explicable (XAI) con Deep Learning, utilizando modelos como LSTM, SVM, DBN y CNN, lo que permite una alta precisión en tiempo real [12, 13, 14, 15, 16].

3. Malware (troyanos, gusanos, ransomware, spyware): El malware incluye una variedad de programas maliciosos que comprometen la seguridad del sistema, tales como troyanos, gusanos, ransomware y spyware. La IA explicable (XAI) ayuda a mejorar la detección en tiempo real utilizando técnicas de ML y DL. El método RMED, con un clasificador de ML y características de encabezados de archivos PE, logra detectar malware con una precisión del 98.42% [17, 18, 19, 20, 21].

4. Amenazas cibernéticas avanzadas (malware, phishing, DoS, APT, intrusiones): Este tipo de ataque incluye un conjunto más sofisticado de técnicas, como malware, phishing, DoS (Denegación de Servicio), APT (Amenazas Persistentes Avanzadas) e intrusiones en sistemas. Para enfrentar estos ataques complejos, se emplean técnicas de Deep Learning y Reinforcement Learning para detectar, simular y responder automáticamente a amenazas en entornos dinámicos. Además, IA Generativa permite tomar decisiones autónomas, lo que mejora la protección de los dispositivos y permite clasificar ataques de phishing con alta precisión [22, 23, 24, 25, 26].

5. SQL Injection: En los ataques de SQL Injection, los atacantes manipulan las consultas SQL utilizadas para acceder o alterar bases de datos sin autorización Para prevenir este tipo de ataque, se emplea GenSQLi, una herramienta que utiliza IA Generativa para crear y probar estos ataques, junto con algoritmos de aprendizaje supervisado y no supervisado para identificar patrones inusuales en las consultas SQL [27, 28].

6. Ataques a datos cifrados (Criptoataques): Los criptoataques buscan explotar debilidades en los sistemas de cifrado, con técnicas como fuerza bruta o explotación de claves débiles. Para mejorar la protección de los sistemas, se utilizan IA y modelos de Machine Learning que comparan algoritmos criptográficos, evaluando su seguridad frente a ataques avanzados y fortaleciendo el cifrado [29].

7. Amenazas internas: Los ataques provenientes de empleados o personal autorizado son difíciles de detectar debido al acceso legítimo que tienen a los sistemas. La implementación de IA y aprendizaje automático permite analizar patrones de comportamiento utilizando modelos predictivos y minería de datos, lo que ayuda a anticipar y prevenir amenazas antes de que ocurran [30].

8. Ataques maliciosos sobre DNS over HTTPS (DoH): Los ataques DoH aprovechan el cifrado de tráfico DNS sobre HTTPS para ocultar actividades maliciosas como exfiltración de datos o malware. Para detectar y protegerse de estos ataques, se aplica un modelo híbrido con Random Forest (RF) y Adaboost (ADT), lo que optimiza la precisión y disminuye el tiempo de procesamiento mediante reducción de características y balanceo de clases [31].

Con esto, podemos ver cómo la inteligencia artificial no solo optimiza la identificación de amenazas en tiempo real, sino que también mejora la respuesta frente a los ataques, adaptándose a nuevas amenazas y mejorando la efectividad en la protección de los sistemas. La Tabla VI resume los métodos impulsados por IA para detectar y mitigar amenazas según el tipo de ataque.

Tipo de Ataque	Definición	Método de detección y prevención de vulnerabilidades	Referencia
Dominio Malicioso	Dominios web utilizados para phishing, distribución de malware o control de bots, generados por algoritmos con el fin de evitar ser detectados	Se utiliza Red Neuronal Superficial (SNN) para mejorar la precisión y el rendimiento en la detección de dominios maliciosos en tiempo real	[11]
Ataques en IoT (Intrusión)	Ataques cibernéticos dirigidos a redes IoT, como intrusiones y explotación de vulnerabilidades	Se usa IA explicable (XAI) con Deep Learning usando modelos LSTM, SVM, DBN y CNN para detectar y responder a ciberataques en tiempo real con alta precisión.	[12, 13,14, 15, 16]
Malware (troyanos, gusanos, ransomware, spyware)	Tipos de malware que comprometen la seguridad del sistema a través de infección, robo de datos o bloqueo de funciones	Se utiliza técnicas de IA explicable (XAI) usando modelos como Machine Learning y Deep Learning, que ayudar a mejorar la detección de amenazas en tiempo real. Para malware en archivos ejecutables, se propone el método RMED, este utiliza un clasificador de ML con features de PE headers para detectar malware en tiempo real, alcanzando una precisión del 98.42%	[17, 18. 19, 20, 21]
Amenazas cibernéticas avanzadas (malware, phishing, DoS, APT, intrusiones)	Ataques como malware, phishing, escaneo de puertos, denegación de servicio (DoS), amenazas persistentes avanzadas (APT) e intrusiones en sistemas	Se emplea Deep Learning y Reinforcement Learning para detectar, simular y responder automáticamente a amenazas en entornos dinámicos, mejorando la precisión y adaptabilidad defensiva. Otra opción es usar aprendizaje automático (ML) para predecir amenazas, aplicar análisis predictivo y automatizar respuestas mediante detección de anomalías y redes neuronales. Tambien usando IA Generativa para tomar decisiones autónomas para proteger los dispositivos. Otra forma, sería una metodología basada en IA y método de ensamblaje para detectar y clasificar phishing de diferentes tipos con alta precisión	[22, 23, 24, 25, 26]
SQL Injection	Ataque que consiste en manipulas consultas SQL mal hechas, para acceder o modificar en la base de datos	Se utiliza la herramienta GenSQLi, que utiliza IA Generativa para crear y probar ataques de SQL Injection, permitiendo generar reglas adaptativas los WAFs. Tambien se usa algoritmos de aprendizaje supervisado y no supervisado para detectar patrones inusuales en las consultas SQL	[27, 28]
Ataques a datos cifrados (Criptoataques)	Ataques que intentan vulnerar sistemas de cifrado mediante técnicas como ataques de fuerza bruta o explotación de claves débiles	El enfoque que utiliza IA y modelos de ML, que comparan algoritmos simétricos y asimétricos que evalúan la eficiencia y seguridad frente a ataques criptográficos avanzados con el objetivo de fortalecerlos.	[29]
Amenazas internas	Ataques ocasionados por empleados o personal autorizado, con el fin de comprometer la seguridad de los datos	Implementación de IA y aprendizaje automático para analizar patrones de comportamiento usando modelos predictivos y minería de datos, permitiendo anticipar y prevenir amenazas antes de que sucedan	[30]
Ataques maliciosos sobre DNS over HTTPS (DoH)	Uso de DoH para ocultar tráfico malicioso como malware o exfiltración de datos al eludir filtros tradicionales de DNS.	Se aplica un modelo híbrido con Random Forest (RF) para identificar tráfico DoH y Adaboost (ADT) para detectar DoH malicioso. Se usa PCA para reducir características y RUS para balancear las clases, logrando alta precisión y bajo tiempo de procesamiento	[31]

TABLA VI
TIPOS DE ATAQUES IMPULSADOS POR IA Y VULNERABILIDADES EN LA CIBERSEGURIDAD TRADICIONAL

C. ¿Qué tan efectivo son las metodologías tradicionales comparado con el uso de IA para la identificación de patrones y la mitigación de amenazas emergentes?

Los sistemas con IA, como los modelos de aprendizaje automático (ML) y aprendizaje profundo (DL), son capaces de identificar patrones desconocidos, aunque gestionar grandes volúmenes de datos puede resultar en tiempos de procesamiento más extensos [32].

En lugar de una intervención manual, la IA permite la automatización. Sin embargo, los modelos de lenguaje a gran escala (LLMs) aplicados en IA son mucho más efectivos, detectando patrones sutiles y ataques que los métodos convencionales no pueden reconocer [33].

Para las Amenazas Persistentes Avanzadas (APT), los sistemas tradicionales a menudo fallan debido a la capacidad de los atacantes para permanecer ocultos durante largos periodos. Los sistemas con IA, por su parte, logran identificar patrones anómalos y correlacionar datos a lo largo del tiempo, mejorando la detección. El uso de redes neuronales gráficas (GNN) ha mostrado ser eficaz en identificar estas amenazas complejas [25].

Los sistemas tradicionales también tienen dificultades para protegerse contra ataques a aplicaciones web como inyecciones SQL y XSS, debido a configuraciones deficientes. La IA mejora la detección de vulnerabilidades al optimizar herramientas de pruebas de penetración automatizadas (VAPT), logrando una mayor precisión y menos falsos positivos [34].

Finalmente, los ransomware suelen eludir los sistemas tradicionales, pero los modelos impulsados por IA, como Random Forest y redes neuronales artificiales, ofrecen una detección mucho más precisa, alcanzando una efectividad de hasta el 99.84% [35].

Con esto podemos decir, que, aunque los sistemas tradicionales sigan siendo útiles para detectar amenazas conocidas, la IA proporciona una ventaja clara al mejorar la capacidad de detectar patrones nuevos y sofisticados, adaptándose con las nuevas amenazas.

La Tabla VII compara la efectividad de los sistemas de seguridad tradicionales con las soluciones basadas en IA. Se muestra cómo los métodos convencionales son vulnerables a ataques sofisticados, mientras que la IA (a través de Machine Learning y LLMs) mejora la detección, ofrece mayor precisión y se adapta mejor a amenazas complejas como intrusiones, phishing y ransomware.

Tipo de Ataques	Efectividad de los sistemas tradicionales	Efectividad de los sistemas con IA	Referencia
Intrusiones (DoS, DDoS, escaneo de puertos)	Solo detectan ataques conocidos, limitando la protección contra ataques nuevos.	Los sistemas con IA (ML/DL) mejoran la detección de patrones desconocidos, lo que mejora la capacidad de detectar ataques. Sin embargo, a grandes cantidades de datos, se vuelve más complicado y require más tiempo de procesamiento para analizarlos	[32]
Phishing generado por IA	Los filtros tradicionales no pueden detectar correos electrónicos generados por IA que son altamente sofisticados y difíciles de identifica	Los LLMs se especializan en la detección de correos electrónicos generados por IA, con la capacidad de identificar patrones sutiles de ataques sofisticados que los métodos convencionales no pueden reconocer	[33]
Advanced Persistent Threats (APT)	Los sistemas tradicionales suelen no detectar APTs por su capacidad para adaptarse y mantenerse invisibles, ya que los ataques se desarrollan a lo largo del tiempo, dificultando su identificación con métodos convencionales.	Los sistemas con IA y aprendizaje automático mejoran la detección al identificar patrones anómalos y correlacionar datos a lo largo del tiempo. El uso de Graph Neural Networks (GNN) ha demostrado ser efectivo para detectar APTs al identificar relaciones complejas durante varias fases del ataque.	[25]
Ataques a Aplicaciones Web (SQL Injection, XSS, CSRF, etc.)	Los sistemas tradicionales de seguridad, como los firewalls o WAFs, pueden ser superados si no se configuran o actualizan adecuadamente, lo que los hace susceptibles a ataques avanzados.	IA mejora la precisión de las herramientas de VAPT, lo que optimiza la identificación de vulnerabilidades de manera más eficiente y con menor tasa de falsos positivos.	[34]
Ransomware (Locker, Crypto)	Los sistemas tradicionales, como los antivirus y firewalls, tienen dificultades para detectar ransomware debido a la criptografía y técnicas avanzadas de evasión.	Los sistemas basados en IA, como Random Forest (RF) y Redes Neuronales Artificiales (ANN), optimizan la detección, logrando una precisión de hasta 96.27% con Random Forest y 99.84% con ANN en la identificación de ransomware.	[35]

TABLA VII
COMPARACIÓN DE LA EFECTIVIDAD DE LOS SISTEMAS TRADICIONALES FRENTE A LA IA EN LA DETECCIÓN DE ATAQUES

D. ¿Cuál es el efecto de la inteligencia artificial en la tasa de detección, la reducción de tiempo para mitigar ataques y la mejora en la seguridad?

La inteligencia artificial (IA) ha demostrado ser crucial en la mejora de la detección de amenazas, la reducción del tiempo necesario para mitigar ataques y el aumento de la seguridad general. Los sistemas impulsados por IA son mucho más eficientes en comparación con los métodos tradicionales, proporcionando una respuesta más rápida y precisa ante ciberataques.

En el caso del ransomware, los sistemas con IA, como las redes neuronales, pueden reducir la mitigación de estos ataques en hasta un 90% al automatizar tareas como la actualización de reglas de firewall y el aislamiento rápido de archivos maliciosos. Esto permite una reacción más ágil y precisa, mejorando considerablemente la seguridad [36, 37].

Para el malware, las redes neuronales profundas (DNN) han mostrado ser extremadamente efectivas, identificando amenazas con una precisión del 99%. Esto no solo mejora la tasa de detección, sino que también reduce el tiempo de mitigación en un 90% en comparación con los enfoques tradicionales, permitiendo una respuesta más rápida y precisa ante las amenazas [37, 38].

En cuanto a la explotación de vulnerabilidades, los sistemas impulsados por IA son capaces de detectar y mitigar vulnerabilidades en cuestión de segundos, lo que mejora la tasa de detección de hasta un 95%. Este nivel de eficiencia es vital para prevenir ataques que explotan fallos de seguridad no identificados previamente [38].

Finalmente, la detección de ataques DoS y DDoS mejora significativamente con el uso de IA, reduciendo el tiempo de respuesta en un 80%. Esto permite que las organizaciones respondan mucho más rápido a estos ataques, que pueden ser devastadores si no se gestionan adecuadamente [10, 23].

En resumen, la integración de IA en sistemas de seguridad es más eficiente no solo aumentando la tasa de detección de amenazas, sino que también reduce considerablemente el

tiempo para mitigar los ataques, mejorando así la seguridad general de los sistemas.

La Tabla VIII resume la alta efectividad de los sistemas de inteligencia artificial (IA) en la mitigación de diversos tipos de ciberataques. Detalla cómo la IA, a través de redes neuronales y otras técnicas avanzadas, logra reducir significativamente las tasas de mitigación y los tiempos de respuesta en ataques como ransomware, malware, explotación de vulnerabilidades y ataques DoS/DDoS, superando ampliamente a los métodos tradicionales.

TABLA VIII
TABLA VI. EFECTIVIDAD DE LA INTELIGENCIA ARTIFICIAL EN LA MITIGACIÓN DE CIBERATAQUES

Tipo de ataques	Tasa de reducción de mitigación de Sistemas impulsados por IA	Referencia
Ransomware	Los sistemas con IA, como las redes neuronales, pueden reducir la mitigación de ataques de ransomware hasta en un 90% al automatizar respuestas, como actualizar reglas de firewall y aislar archivos maliciosos rápidamente.	[36, 37]
Malware	Los sistemas de IA, como las redes neuronales profundas (DNN), identifican malware con un 99% de precisión, reduciendo el tiempo de mitigación en un 90% en comparación con los métodos tradicionales.	[37, 38]
Explotación de vulnerabilidades	Los sistemas con IA detectan y mitigan vulnerabilidades en cuestión de segundos, mejorando la tasa de detección de hasta un 95%	[38]
Dos y DDoS	La detección de ataques DoS y DDoS mejora significativamente con el uso de IA, reduciendo el tiempo de respuesta en un 80%	[10, 23]

E: Análisis de la Evidencia Práctica:

Los sistemas de Darktrace demuestran cómo la IA puede aprender el "comportamiento normal" de una red para identificar anomalías en tiempo real, detectando amenazas sofisticadas que un firewall tradicional no podría reconocer.

La tecnología de Fortinet ejemplifica esto al responder de forma autónoma a ataques de día cero, aislando el malware en segundos, lo que reduce el tiempo de mitigación y el daño potencial.

IV. DISCUCIÓN

La inteligencia artificial (IA) está cambiando por completo la forma en que las empresas protegen su información. A diferencia de los métodos tradicionales, que funcionan con reglas fijas y detectan amenazas conocidas, la IA tiene la capacidad de aprender, adaptarse y reconocer comportamientos sospechosos que antes pasaban desapercibidos. Gracias a esto, se pueden identificar ataques nuevos, responder más rápido y prevenir muchos problemas antes de que causen daño. Por eso, cada vez más organizaciones están invirtiendo en herramientas basadas en IA y formando a su personal para estar mejor preparados [32].

Sin embargo, también hay que tener en cuenta que los atacantes están utilizando la misma tecnología para mejorar sus ataques. Usan la IA para esconderse entre el tráfico normal [4], controlar sistemas desde plataformas legítimas como Dropbox [5], y crear mensajes falsos muy realistas con ayuda de modelos como ChatGPT [6][7][8]. Incluso algunos ataques aprendan de sus errores y cambian su comportamiento para evitar ser bloqueados [9][10]. Esto hace que los sistemas antiguos, por sí solos, ya no sean suficientes para proteger a las empresas frente a estas amenazas más inteligentes.

Frente a eso, la IA se convierte en una gran aliada. Permite detectar sitios peligrosos en internet [11], analizar redes con muchos dispositivos conectados [12][13] o encontrar virus y programas maliciosos con mucha precisión [17][18][19]. En situaciones más complejas, como los ataques avanzados o múltiples a la vez, puede dar una respuesta rápida y automática [22][24][24]. También ayuda a detectar comportamientos extraños dentro de la empresa o cuando se usan conexiones cifradas para ocultar actividades [30][31].

A comparación de los métodos tradicionales, la IA ofrece muchas ventajas. Por ejemplo, puede identificar correos falsos generados por otras IA, algo que los filtros comunes no logran detectar [33]. También es capaz de notar amenazas que se mantienen escondidas por mucho tiempo [25] y mejora las pruebas de seguridad en páginas web o bases de datos [34]. Incluso frente a ataques como el ransomware, la precisión puede llegar al 99.84% [35].

Además de ser más precisa, la IA también es más rápida. Muchos ataques que antes requerían horas para ser controlados ahora pueden ser detenidos en minutos o incluso segundos, lo que es vital para evitar pérdidas o daños mayores [36][37][38].

Pero a pesar de todos estos beneficios, también existen algunas limitaciones y riesgos que deben tenerse en cuenta. Por ejemplo, aunque la IA puede procesar grandes cantidades de información, esto requiere un alto poder de cómputo, lo que implica mayor consumo de energía y más costos para las empresas. También hay un riesgo si los sistemas de IA son manipulados o si dependen demasiado de datos poco confiables, lo que puede comprometer su eficacia.

V. CONCLUSIONES

Esta investigación demuestra que la inteligencia artificial (IA) es clave para la ciberseguridad, mejorando la detección y respuesta a amenazas complejas. Casos reales como los de Darktrace y Fortinet confirman que la IA puede identificar anomalías y automatizar respuestas de forma altamente eficiente, superando significativamente a los métodos tradicionales.

A pesar de sus beneficios, la implementación de la IA tiene desafíos, como el alto costo computacional y la necesidad de datos de calidad. Por ello, la IA debe ser vista como una herramienta estratégica que complementa las defensas existentes, no como una solución única. En el futuro, será importante explorar cómo hacer esta tecnología más accesible y transparente para todas las organizaciones.

Se sugiere que las empresas inviertan en la formación de sus equipos de seguridad para que puedan colaborar con las herramientas de IA, entendiendo sus resultados y complementando sus funciones.

Se propone un enfoque estratégico donde la IA sea una herramienta que complemente a los sistemas tradicionales, utilizando **IA explicable (XAI)** para justificar las decisiones del modelo y reducir errores, como los falsos positivos.

VI. AGRADECIMIENTO

Agradezco a la Universidad Tecnológica del Perú por brindar el acceso a recursos académicos que permitieron el desarrollo de esta investigación. También extiendo mi reconocimiento a los docentes del curso Formación para la Investigación Sistemas, quienes orientaron este trabajo con sus recomendaciones y observaciones. Finalmente, agradezco a mis compañeros y familiares por su apoyo constante durante la realización de este estudio.

REFERENCIAS

[1] M. Alanezi y R. M. A. Al-Azzawi, "AI-Powered Cyber Threats: A Systematic Review," Deleted Journal, vol. 4, no. 3, pp. 166-188, 2024, doi: 10.58496/mjcs/2024/021.

[2] E. J. C. Reyes y R. F. B. Medina, "Estructuras metodológicas PICO y PRISMA 2020 en la elaboración de artículos de revisión sistemática: Lo que todo investigador debe conocer y dominar," Ciencia Latina Revista Científica Multidisciplinar, vol. 9, no. 1, pp. 8525-8543, 2025, doi: 10.37811/cl_rcm.v9i1.16491.

[3] *Elsevier, "Acerca de Scopus | Base de datos de resúmenes y citas," www.elsevier.com. [En línea]. Disponible en: https://www.elsevier.com/es-es/products/scopus?utm_source=chatgpt.com. (Consultado: Ago. 1, 2025).*

[4] L. Coppolino, S. D'Antonio, G. Mazzeo y F. Uccello, "The good, the bad, and the algorithm: The impact of generative AI on cybersecurity," Neurocomputing, vol. 623, 2025, doi: 10.1016/j.neucom.2025.129406.

[5] M. Jeong, J. Park y S. H. Oh, "Cyber Environment Test Framework for Simulating Command and Control Attack Methods with Reinforcement Learning," Appl. Sci. Switzerland, vol. 15, no. 4, 2025, doi: 10.3390/app15042120.

[6] M. Bethany et al., "Lateral Phishing With Large Language Models: A Large Organization Comparative Study," IEEE Access, vol. 13, pp. 60684–60701, 2025, doi: 10.1109/ACCESS.2025.3555500.

[7] E. Iturbe, O. Llorente-Vazquez, A. Rego, E. Rios y N. Toledo, "Unleashing offensive artificial intelligence: Automated attack technique code generation," Comput. Security, vol. 147, 2024, doi: 10.1016/j.cose.2024.104077

[8] [8] M. A. Elsadig, "ChatGPT and Cybersecurity: Risk Knocking the Door," *J. Internet Services and Inform. Security*, vol. 14, no. 1, pp. 1–15, 2024, doi: 10.58346/JISIS.2024.I1.001.

[9] [9] B.-S. Kim *et al.*, "Optimal Cyber Attack Strategy Using Reinforcement Learning Based on Common Vulnerability Scoring System," *CMES Comput. Model. Eng. Sci.*, vol. 141, no. 2, pp. 1551–1574, 2024, doi: 10.32604/cmes.2024.052375.

[10] [10] A. M. AL-Hawamleh, "Predictions of Cybersecurity Experts on Future Cyber-Attacks and Related Cybersecurity Measures," *Int. J. Advanced Comput. Sci. Appl.*, vol. 14, no. 2, pp. 801–809, 2023, doi: 10.14569/IJACSA.2023.0140292.

[11] [11] J. Senanayake *et al.*, "MADONNA: Browser-based malicious domain detection using Optimized Neural Network by leveraging AI and feature analysis," *Comput. Security*, vol. 152, 2025, doi: 10.1016/j.cose.2025.104371.

[12] [12] W. Serrano, "CyberAIBot: Artificial Intelligence in an intrusion detection system for CyberSecurity in the IoT," *Future Gener. Comput. Syst.*, vol. 166, 2025, doi: 10.1016/j.future.2024.107543.

[13] [13] M. Ragab *et al.*, "Artificial intelligence driven cyberattack detection system using integration of deep belief network with convolution neural network on industrial IoT," *Alexandria Eng. J.*, vol. 110, pp. 438–450, 2025, doi: 10.1016/j.aej.2024.10.009.

[14] [14] A. S. R. M. *et al.*, "Explainable artificial intelligence in web phishing classification on secure IoT with cloud-based cyber-physical systems," *Alexandria Eng. J.*, vol. 110.0, pp. 490.0, 2025, doi: 10.1016/j.aej.2024.09.115.

[15] [15] S. B. Sharma y A. K. Bairwa, "Leveraging AI for Intrusion Detection in IoT Ecosystems: A Comprehensive Study," *IEEE Access*, vol. 13, pp. 66290–66317, 2025, doi: 10.1109/ACCESS.2025.3550392.

[16] [16] M. R. WAR, Y. SINGH, Z. A. SHEIKH y P. K. SINGH, "REVIEW ON THE USE OF FEDERATED LEARNING MODELS FOR THE SECURITY OF CYBER-PHYSICAL SYSTEMS," *Scalable Comput.*, vol. 26, no. 1, pp. 16–33, 2025, doi: 10.12694/scpe.v26i1.3438.

[17] [17] A. el Hariri, M. Mouiti y M. Lazaar, "Realtime ransomware process detection using an advanced hybrid approach with machine learning within IoT ecosystems," *Eng. Res. Express*, vol. 7, no. 1, 2025, doi: 10.1088/2631-8695/ada3b3.

[18] [18] H. Manthena *et al.*, "Explainable Artificial Intelligence (XAI) for Malware Analysis: A Survey of Techniques, Applications, and Open Challenges," *IEEE Access*, vol. 13, pp. 61611–61640, 2025, doi: 10.1109/ACCESS.2025.3555926.

[19] [19] A. H. Salem, S. M. Azzam, O. E. Emam y A. A. Abohany, "Advancing cybersecurity: a comprehensive review of AI-driven detection techniques," *J. Big Data*, vol. 11, no. 1, 2024, doi: 10.1186/s40537-024-00957-y.

[20] [20] A. Galli, V. la Gatta, V. Moscato, M. Postiglione y G. Sperlì, "Explainability in AI-based behavioral malware detection systems," *Comput. Security*, vol. 141, 2024, doi: 10.1016/j.cose.2024.103842.

[21] [21] K. Soliman, M. Sobh y A. M. Bahaa-Eldin, "Robust Malicious Executable Detection Using Host-Based Machine Learning Classifier," *Comput. Mater. Continua*, vol. 79, no. 1, pp. 1419–1439, 2024, doi: 10.32604/cmc.2024.048883.

[22] [22] S. Khanzadeh *et al.*, "An exploratory study on domain knowledge infusion in deep learning for automated threat defense," *Int. J. Inform. Security*, vol. 24, no. 1, 2025, doi: 10.1007/s10207-025-00987-4.

[23] [23] M. Khayat *et al.*, "Empowering Security Operation Center with Artificial Intelligence and Machine Learning - A Systematic Literature Review," *IEEE Access*, vol. 13, pp. 19162–19197, 2025, doi: 10.1109/ACCESS.2025.3532951.

[24] [24] S. Gupta y B. Crispo, "Towards autonomous device protection using behavioural profiling and generative artificial intelligence," *IET Cyber-Physical Syst. Theory Appl.*, vol. 10, no. 1, 2025, doi: 10.1049/cps2.12102.

[25] [25] R. Buchta, G. Gkoktsis, F. Heine y C. Kleiner, "Advanced Persistent Threat Attack Detection Systems: A Review of Approaches, Challenges, and Trends," *Digital Threats: Res. Pract.*, vol. 5, no. 4, 2024, doi: 10.1145/3696014.

[26] [26] Y. A. Alsariera, M. H. Alanazi, Y. Said y F. Allan, "An Investigation of AI-based Ensemble Methods for the Detection of Phishing Attacks," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 3, pp. 14266–14274, 2024, doi: 10.48084/etasr.7267.

[27] [27] V. Babaey y A. Ravindran, "GenSQLi: A Generative Artificial Intelligence Framework for Automatically Securing Web Application Firewalls Against Structured Query Language Injection Attacks," *Future Internet*, vol. 17, no. 1, 2025, doi: 10.3390/fi17010008.

[28] [28] N. Augustine, A. B. M. Sultan, M. H. Osman y K. Y. Sharif, "Application of Artificial Intelligence in Detecting SQL Injection Attacks," *Int. J. Inform. Visualiz.*, vol. 8, no. 4, pp. 2131–2138, 2024, doi: 10.62527/joiv.8.4.3631.

[29] [29] N. Kshetri *et al.*, "algoTRIC: Symmetric and Asymmetric Encryption Algorithms for Cryptography – A Comparative Analysis in AI Era," *Int. J. Advanced Comput. Sci. Appl.*, vol. 15, no. 12, pp. 1–14, 2024, doi: 10.14569/IJACSA.2024.0151201.

[30] [30] E. Yilmaz y O. Can, "Unveiling Shadows: Harnessing Artificial Intelligence for Insider Threat Detection," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 2, pp. 13341–13346, 2024, doi: 10.48084/etasr.6911.

[31] [31] Q. Abu Al-Haija, M. Alohaly y A. Odeh, "A Lightweight Double-Stage Scheme to Identify Malicious DNS over HTTPS Traffic Using a Hybrid Learning Approach," *Sensors*, vol. 23, no. 7, 2023, doi: 10.3390/s23073489.

[32] [32] M. A. Umar, Z. Chen, K. Shuaib y Y. Liu, "Effects of feature selection and normalization on network intrusion detection," *Data Sci. Manag.*, vol. 8, no. 1, pp. 23–39, 2025, doi: 10.1016/j.dsm.2024.08.001.

[33] [33] J. Zhang, P. Wu, J. London y D. Tenney, "Benchmarking and Evaluating Large Language Models in Phishing Detection for Small and Midsize Enterprises: A Comprehensive Analysis," *IEEE Access*, vol. 13, pp. 28335–28352, 2025, doi: 10.1109/ACCESS.2025.3540075.

[34] [34] A. Alquwayzani, R. Aldossri y M. Frikha, "Mitigating Security Risks in Firewalls and Web Applications using Vulnerability Assessment and Penetration Testing (VAPT)," *Int. J. Advanced Comput. Sci. Appl.*, vol. 15, no. 5, pp. 1348–1364, 2024, doi: 10.14569/IJACSA.2024.01505136.

[35] [35] D. Smith, S. Khorsandroo y K. Roy, "Machine Learning Algorithms and Frameworks in Ransomware Detection," *IEEE Access*, vol. 10, pp. 117597–117610, 2022, doi: 10.1109/ACCESS.2022.3218779.

[36] [36] A. al Siam *et al.*, "A Comprehensive Review of AI's Current Impact and Future Prospects in Cybersecurity," *IEEE Access*, vol. 13, pp. 14029–14050, 2025, doi: 10.1109/ACCESS.2025.3528114.

[37] [37] D. A. I.M. *et al.*, "Artificial Intelligence-Based Malware Detection, Analysis, and Mitigation," *Symmetry*, vol. 15, no. 3, 2023, doi: 10.3390/sym15030677.

[38] [38] M. Ozkan-Okay *et al.*, "A Comprehensive Survey: Evaluating the Efficiency of Artificial Intelligence and Machine Learning Techniques on Cyber Security Solutions," *IEEE Access*, vol. 12, pp. 12229–12256, 2024, doi: 10.1109/ACCESS.2024.3355547.